## Time, Money, and Effort: A Practical Approach to Digital Content Management

Christine Wiseman and Al Matthews

### Introduction

For libraries and archives, the digital content management and preservation landscape is rapidly evolving. As digital projects evolve into digital programs focusing on the mass digitization of entire collections, institutions are faced with ensuring long term accessibility to vast quantities of digital assets. "Most institutions," according to a Portico and Cornell University Library report, "are only beginning to understand that their investment in creating digital collections must be met with a commitment and infrastructure to protect this content for its lifetime." As digital collections grow exponentially, institutions are faced with the challenge of providing continued access as well as long term preservation. The systems and options for the management, presentation, and preservation of digital assets are numerous. Each has its pros and cons, whether an out-of-the-box, vendor-provisioned system, or an open-source application where the source code is free and openly available for use and modification. Some platforms focus on preservation, others on presentation, and still others on content management. Company mergers, upgrades – and even dissolutions – further complicate the problem. Like many mid-sized academic institutions, the Atlanta University Center Robert W. Woodruff Library (AUC Woodruff Library) found itself in need of assessment and consolidation of existing digital content management platforms as digital collections rapidly expand. This article addresses the process this institution undertook to evaluate the digital content management and preservation landscape in order to inform future growth and expansion of its digital program.

### Background

Established in 1982, the AUC Woodruff Library is unique on a number of fronts. It is an independent, non-profit academic library and research center providing information services to the world’s

largest consortium of Historically Black Colleges and Universities (HBCUs): Clark Atlanta University, the Interdenominational Theological Center, Morehouse College and Spelman College. The Archives Research Center's (ARC) history dates back to the establishment of the Collection in 1925 under the auspices of Atlanta University's Trevor Arnett Library "Negro Collection." The archival program and collections were transferred to AUC Woodruff Library upon its establishment as the Library for the Atlanta University Center (AUC) schools in 1982. Guided by its 2010-2015 Strategic Plan (Building a 21st Century Learning Community – Advancing the Academic Village), the AUC Woodruff Library's mission is to serve as the center of the academic village for its member institutions, providing the highest level of information resources and services in support of teaching and learning, scholarship and cultural preservation of the Atlanta University Center. Expanding access to digital collections and building a preservation program for the collection in all formats are primary objectives in the Library's strategic plan.

In a recent OCLC survey, 97% of 169 libraries surveyed have completed at least one digitization project and/or have an active program in place. In step with this trend, the AUC Woodruff Library has nearly ten years of experience developing digital services, programs, and collections that expand access to hidden primary-resource collections. Depending on the size, format, and complexity of the project, the library engages in both in-house digital conversion and outsourcing to vendors. Adherence to professional standards is a primary objective in all digital initiatives.

Digital initiatives at the AUC Woodruff Library date back to 2005, beginning with a five year partnership with Cornell University, SOLINET (now LYRASIS), and nine historically black colleges and universities to expand access to the founding documents of HBCUs. As a result of this partnership, the AUC Woodruff Library gained expertise in digitization standards, metadata creation, and digitization methodologies. The AUC Woodruff Library serves as the technical administrator of the HBCU Library Alliance Digital Collection (http://hbcudigitallibrary.auctr.edu/), a collection that has grown from approximately 9,000 to more than 16,000 images representing the founding materials of 22 HBCUs. Images of materials dating from the early 1800s to the present document the role of HBCUs in

the history of African-American higher education. In 2006, the Library became the custodian for the Morehouse College Martin Luther King Jr. Collection, a collection of 10,000 of his personal items. As a result of this custodianship, the Library has engaged in processing the collection, creating item level, web-based, publicly accessible inventories, and digitizing the collection. Dissemination of the collection has been achieved through the web-based inventories, scholarly forums, inclusion in curriculum and instruction, and presentations and articles to professional communities.

Since 2011, the HBCU Library Alliance has preserved digital content in the MetaArchive Cooperative, a distributed LOCKSS ("Lots of Copies Keeps Stuff Safe") digital preservation network for the content held in the HBCU Library Alliance (HBCU LA) Digital Collection. On behalf of the HBCU Library Alliance, The AUC Woodruff Library provides technical support and hosts the server as a preservation node on the network. The HBCU Library Alliance is an active member of the MetaArchive Cooperative; AUC Woodruff Library staff coordinates the ingest of master files and metadata for long term preservation purposes. Staff also participates in monthly conference calls and attends annual meetings. The AUC Woodruff Library uses both CONTENTdm and DigitalCommons, a hosted institutional repository for discovery and access to digital content. The WorldCat Local discovery tool provides access to content in both of these repositories through a "Google-like" search box on the Library's website (see http://www.auctr.edu). In addition, discovery of content within these repositories is also possible using search engines such as Yahoo and Google. Alternatively, a user can navigate directly to either CONTENTdm or DigitalCommons to conduct individual searches, or directly link to digitized content in CONTENTdm through archival finding aids that are searchable via XTF. To further expand access, metadata from the repositories is harvested and pulled into several statewide repositories, including the Digital Library of Georgia and the Georgia Knowledge Repository. Currently there are approximately 77,000 images available in CONTENTdm and over 3,000 publications in DigitalCommons. Most of the accessible digital content consists of still images, manuscripts, and publications such as theses and dissertations; however, the library is greatly expanding the digitization of audio and video collections.

**Literature Review**

While there are a number of existing comparative reports about content management systems in libraries and archives, nothing close to hand was found to be at once current, comprehensive, and applicable to mid-sized academic institutions such as the AUC Woodruff Library. Upon review of reports published over the past decade, Jody DeRidder's 2004 article clearly elucidates the broad landscape of institutional repository software which includes some content management systems still in use. More recently, in 2009, Marill and Luczak of the National Library of Medicine took on a similar investigation of digital repository software at a large government institution with significant information technology infrastructure. Their evaluation of open source and commercial options, list of criteria, and process of narrowing down an initial group of ten systems to three for extensive review struck us as replicable for smaller organizations. In the end, the working group at the NLM recommended building a pilot Fedora repository that would in turn be subject to further testing and consideration.

There is also much to be gained by reviewing the process that a large academic library, such as Yale University, undertook in evaluating digital content management and preservation systems. Yale embarked on a large scale initiative to "create a unified Hydra/Fedora infrastructure for the preservation and dissemination of digital materials through a single search box in Blacklight." From Yale's experience, smaller institutions can look toward this project as a model and choose applicable portions. Fedora/Hydra is scalable and can be moved in either direction: larger or smaller. In its report, Yale justified a significant internal resource allocation, assuming a much broader and more deeply resourced technical infrastructure than our own, but reaching similar conclusions.

Equally valuable is the experience of the Low Country Digital Library's (LCDL) search for a more suitable digital asset management system and their eventual decision to move to an open source option. Heather Gilbert and Tyler Mobley recount moving a consortial digital library from a vendor platform to building an open source solution with just two full-time staff members, neither trained as software developers, although both had significant technical skills. In "Breaking up with CONTENTdm: Why and How One Institution

Took the Leap to Open Source," the authors describe their search for a more scalable platform which would offer improved searching, and the ability to customize the user interface for its project partners. Although the migration proved challenging and they encountered some technical roadblocks, they built a digital library that meets all of their needs. In the end, the LCDL used four open source products: Fedora Commons, Drupal, Backlight and RUcore.

In terms of useful survey models, Hoe-Lian Goh Dion, Alton Chua, et. al. present a comprehensive and simple checklist for the evaluation of open source digital library software, although not all of the features proved relevant to our needs. Split into 12 categories of functional requirements including content management, acquisition, document formats, version control, metadata, privacy and other measures, this resource of enumerated features can be easily adapted. Providing an objective measure of functional requirements is useful; though, in reality an institution must judge carefully the system that best matches their individual needs (keeping in mind that the checklist can be tweaked to address local priorities). Other beneficial resources for smaller and mid-sized institutions include the POWRR (Preserving Digital Objects with Restricted Resources) online portal and tool evaluation grid, and the University of Toronto Libraries' poster depicting their migration from CONTENTdm to Islandora. To complement the formal literature review, library staff also engages with local intuitions on a more informal basis to share information and experiences related to digital content management issues. Specifically, AUC Woodruff Library staff participates in the Atlanta Area Digital Archivists, an Atlanta area group of professionals from area institutions including Georgia Tech, Emory University, and Georgia State University, that meets quarterly to discuss common issues related to digital preservation, digital curation and content management. Formed in 2013, the group shares best practices and documentation and seeks opportunities for shared training and other areas of potential collaboration.

**Methodology**

Within the AUC Woodruff Library, the Digital Services Unit (DSU) is responsible for managing and implementing digital conversion projects, providing access to digital content, as well as library systems administration. The DSU is comprised of a Unit

Head, a Systems Librarian, a Bibliographic Services Librarian, a Software Developer with systems administrative responsibilities, and a Scanning Technician. Situated within the larger context of the Content and Collections Management Department, the DSU was formed in 2008 to directly support the library's strategic goals related to expanding access to digital content and preservation of collections in all formats.

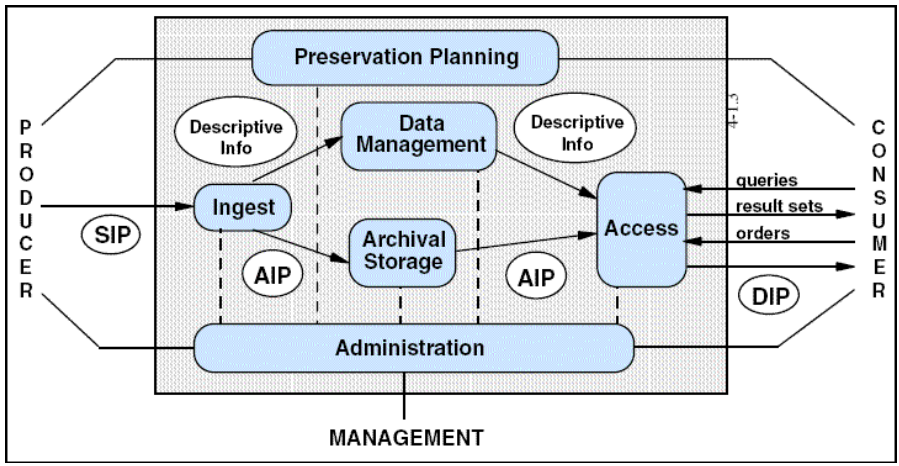
DSU members work collaboratively with the Archives Research Center staff because the bulk of the digital content originates from their collections; a premier archives of primary source materials on African American and African Diaspora history and culture. In 2013, the library formed the Content Management Evaluation Working Group (CMEWG) to issue recommendations regarding digital collections storage, management, and preservation technology. The CMEWG is comprised of members from the Library's Digital Services Unit including the Unit Head, Software Development Specialist, and Library Technical Assistant. The working group consulted with other staff members as needed, including staff from information technology and archives. The overarching goal of the CMEWG was to implement a holistic approach to planning for the conversion, storage, preservation, and access of digital collections in all formats. The CMEWG met monthly over about six months to review the literature, select systems for consideration, and determine criteria and process for evaluation. In the process of evaluating digital content management systems, the working group considered a myriad of features. Also taken into account were the following overarching goals to provide context and focus throughout the process:

- consolidation of digital collections
- interoperability with existing library systems
- digital asset management functionality for master and access files
- long term preservation of digital assets.

A primary consideration during the evaluation was where each system fit within the Open Archival Information System (OAIS) Reference Model (see Table 1). The OAIS Reference Model, developed by the Consultative Committee for Space Data Systems, provides a functional framework for what is required in a repository to preserve and provide access to digital content over the long term.

There are five functional entities: ingest, archival storage, data management, administration, and access. An ISO standard (ISO 14721:2003), the OAIS Reference Model is widely adopted as the standard model for developing digital preservation systems (http://www.iso.org/iso/catalogue_detail.htm?csnumber=24683).

**Table 1: OAIS Reference Model**




Source: Reference Model for an Open Archival Information System (OAIS); Consultative Committee for Space Data System, Washington, DC, 2001; 4:1. http://public.ccsds.org/publications/archive/650x0b1s.pdf.

Another area of consideration for the CMEWG is the AUC Woodruff Library's focus on the digitization of original audio and video resources. The preservation of audiovisual formats, especially obsolete magnetic formats, is designated as a priority due to the short life expectancy and the obsolescence of playback technology. Digital reformatting of machine dependent audiovisual formats is considered the only option for preservation of these materials. Managing digitized audiovisual materials requires staff expertise, specific workflows, and specialized tools for providing access, all of which present challenges to many institutions. Moreover, uncompressed audiovisual digital masters, as recommended by the American Library Association, result in huge files requiring substantial digital storage, especially when backups are taken into account.

After conducting an extensive literature review, the working group assembled a lengthy list of functional requirements to compare the offerings of various systems. The list of requirements (see Appendix A) was developed from internal brainstorming coupled with examples found in the literature. Requirements were grouped by areas of functionality such as general considerations, formats supported, metadata, access and privacy and preservation features. The CMEWG looked at both vendor and open source products. Vendor products are appealing because they offer out-of-the box functionality and built-in technical support. Drawbacks to vendor products include high licensing fees, lack of ability for customization, and questions of sustainability if the vendor goes out of business. Adopting open source applications, however, may require significant development time as well as necessitate expertise in computer systems and programming. Many open source products are well supported by a network of developers, but documentation and backing can vary widely according to project and popularity. While open source applications are "free" there can be significant costs involved. Because sustainability is a concern with open source software for libraries, another trend is for a third party to serve as the home organization for an application. This may include a development staff, support services, documentation, training, and hosting. Of course, these options typically come with a fee, but institutions gain the flexibility to pay for the level of services needed based on local IT infrastructure and in-house expertise. One fact, quickly realized, was difficulty to evaluate any given system without actually running it. The systems initially considered are listed in Table 2.

**Table 2: Systems Evaluated, grouped by mode of ownership**

| Vendor | Open-Source | Hybrid |
|---|---|---|
| CONTENTdm (OCLC) | Archivematica | DuraCloud |
| DigiTool / Rosetta (Ex Libris) | ArchivesSpace | Fedora 3-Islandora |
| | DSpace | LOCKSS-MetaArchive |
| | Fedora 4-Hydra | |
| | LOCKSS | |
| | Omeka | |

CONTENTdm, Fedora-Hydra, and Fedora-Islandora are described in additional detail later in this report. Omeka is a popular publishing platform that is particularly useful for creating online digital exhibits. Once installed onto a library's server, Omeka offers an array of plug-ins and customizations; alternatively, an institution can run the hosted version, Omeka.net. Based at Stanford University, LOCKSS is a specialized tool used in some cases for access to journals, and also serves as a "dark" archive for permanent digital preservation. Libraries can participate in a private LOCKSS network such as the MetaArchive Cooperative. DuraCloud is a competing preservation tier, which now integrates closely with Archivematica. Rosetta is ExLibris' new digital preservation system, integrating with ExLibris' DigiTool repository. Archivematica closely follows the OAIS reference model, as a loosely integrated collection of independent programs or "microservices" that focuses on the automation of various workflows and processes from producer, through archive, to consumer.

In the evaluation, the working group noticed that the functionality of content management systems could be grouped into three broad categories: 1) presentation of digital surrogates; 2)

repository management; and 3) asset preservation. To better understand how each system represents these categories of functionality, the CMEWG developed a graphical matrix shown in Table 3. For each system the categories of functionality are rated in a scale of 0-3 in ascending order of focus. For example, LOCKSS is used for preservation, yet offers minimal front end searching and display (e.g. presentation). Omeka's concentration is in presenting information online while preservation is not the focus. Cost and time are rated on a two point scale which compares the difference between licensing a vendor product that may require few in-house resources to operate, versus adopting a low cost open source system that necessitates a great deal of in-house development. Because these products are changing rapidly, the column on the far right includes a brief summary of relevant updates.

**Table 3: Systems Matrix**

| | | | | | |
|---|---|---|---|---|---|
| Hydra | ■■■<br>■■■<br>■ | presentation<br>repository<br>preservation | ■<br>■■ | cost<br>time | Designed to work with a front end for Fedora 4 repositories. New IMLS funded initiative to develop an out of the box solution. |
| Islandora | ■■■<br>■■■<br>■ | presentation<br>repository<br>preservation | ■<br>■■ | cost<br>time | Islandora is a Drupal-friendly front end to Fedora 3. LYRASIS has used it to build a repository with hosting. |
| Duracloud | ■■<br>■■■ | presentation<br>repository<br>preservation | ■<br>■ | cost<br>time | A managed, cloud based service from DuraSpace for preserving |

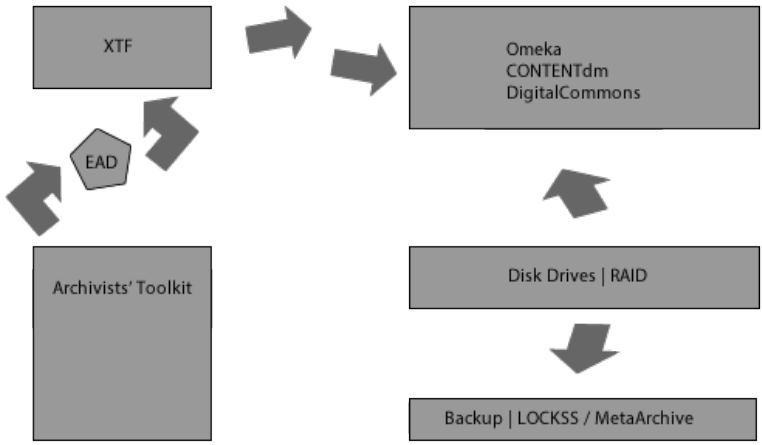
| | | | | | |
|---|---|---|---|---|---|
| | | | | | digital content. Provides back up, syncing, and integrity checking. |
| DSpace | ■■<br>■■<br>■■■ | presentation<br>repository<br>preservation | ■<br>■ | cost<br>time | Intended to be a turnkey institutional repository tool. |
| CONTENTdm | ■■■<br>■■ | presentation<br>repository<br>preservation | ■■<br>■ | cost<br>time | OCLC has not released a preservation repository product. |
| DigiTool | ■■■<br>■■ | presentation<br>repository<br>preservation | ■■<br>■ | cost<br>time | Produced by ExLibris; designed to manage and provide online access to digital assets that compliments Rosetta. |
| Rosetta | ■■<br>■■■ | presentation<br>repository<br>preservation | ■■<br>■ | cost<br>time | See above. |
| Omeka | ■■■<br>■ | presentation<br>repository<br>preservation | ■<br>■ | cost<br>time | Specializes in online exhibits, numerous plug-ins available. |
| LOCKSS | ■<br>■■■ | presentation<br>repository<br>preservation | ■<br>■ | cost<br>time | Dark preservation tool. |

**Systems Architecture: Current and Future**

Considering the pros and cons of each system is critical to the planning process, but the evaluation must be based upon how each system would integrate with an institution's workflows and system

architecture. Table 4 offers a visual representation of the current content management and underlying systems architecture at AUC Woodruff Library. Presently, digital collections are stored and made accessible on three different platforms: two instances of CONTENTdm, the institutional repository DigitalCommons, and Omeka. Only the HBCU Library Alliance collections are preserved in the MetaArchives LOCKSS network. CONTENTdm houses images from the Morehouse College Martin Luther King Jr. Collection, Tupac Amaru Shakur Collection, and the HBCU Library Alliance Collections. DigitalCommons houses theses, dissertations, publications, yearbooks, and video collections. Omeka is used for online digital exhibits.

**Table 4: Current Systems Architecture**




The architecture from an Archives standpoint is based around Archivists' Toolkit Version 2.0.14. Encoded Archival Description (EAD) records are exported from Archivists' Toolkit and presented to XTF 3.0, which presents the finding aid online for web-based searching. Developed and maintained by the California Digital Library, XTF (eXtensible Text Framework) is an open source platform for providing access to digital content. Archivists' Toolkit is now part of a new platform, ArchivesSpace, and the library recently completed this migration. ArchivesSpace offers the option

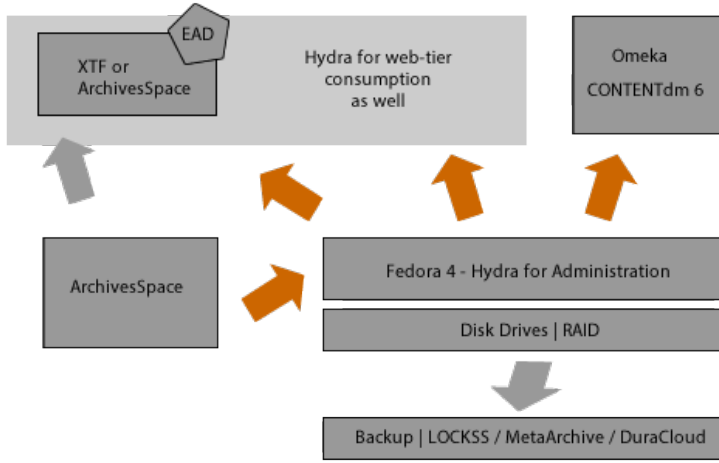
to link to Digital Access Objects, which at present are maintained in CONTENTdm and Omeka. All digital assets, including masters, web images, and associated metadata are stored in the Library's file system using standardized file naming and a system of folders. The file system is backed up offsite for disaster recovery purposes. Metadata follows the Dublin Core standard and is stored in Excel spreadsheets, which are imported or occasionally re-keyed into the applicable content management system during upload of access images.

**Future Architecture**

The addition of a repository layer using Fedora-Hydra in the proposed future architecture (see Table 5) offers a number of possible efficiencies, as well as more robust management of master files and metadata. In this scenario, CONTENTdm 7 and DigitalCommons may become hosted while the Fedora Repository, Archives Space, Omeka and CONTENTdm 6 (MCMLK/TAS) are hosted on library servers.

Migration from Archivists' Toolkit 2.0.14 to ArchivesSpace offers the possibility of deploying the public interface as well as using the back-end for cataloging, which has potential to obviate the need to run the separate instance of XTF for online searching and display of finding aids. So far, the AUC Woodruff Library does not anticipate that it will present the public interface in the current version of ArchivesSpace due to somewhat confusing search display results. However, the public ArchivesSpace would offer certain advantages over XTF, especially related to workflow, because it eliminates the need to export the EAD record to XTF.

**Table 5: Systems Architecture – Future**




Fedora is open source software that provides a repository system for the management of digital content that is designed to meet the preservation and access needs of digital libraries and archives. It has a worldwide support base among academic, research, government, and cultural heritage organizations. Serving as the back-end repository, Fedora functions best when integrated with software that excels in presentation and display of digital content. Some of the most common configurations include pairing Fedora with Hydra or Islandora. Desirable features available in Fedora-based repositories (and/or their Hydra or Islandora-based management systems) include some of the following:

| SELECTED HYDRA FEATURES[1] | |
|---|---|
| Multiple file, or folder, upload | User dashboard for file management |
| Flexible user and group-based access controls | User notifications |
| Transcoding of audio and video files | Single-use links |

[1] For the complete list, see https://github.com/projecthydra/sufia/#what-is-sufia, accessed April 27, 2015.

| | |
|---|---|
| Generation and validation of identifiers | Google Analytics for usage statistics |
| Fixity checking | Integration with cloud storage providers |
| Version control | Google Scholar-specific metadata embedding |
| Forms for batch editing metadata | User managed collections for grouping files |
| Faceted search and browse (based on Blacklight) | Full text indexing and searching |
| Social media interaction | Sharing with groups and users |
| User profiles | |

Building a digital repository using Fedora 4 and Hydra would expand capabilities in both management and access of digital content by providing a back-end repository with search and discovery functions for in-house use by staff. A public interface feature could be built-out in a later phase. Although Fedora is not considered a "preservation repository" in our architecture, it does allow for preservation readiness and integration with digital preservation tools such as DuraCloud or MetaArchive (LOCKSS). In addition, a Fedora 4 implementation offers efficiencies and cost savings due to its storage configuration. With version 4, Fedora becomes storage-agnostic, permitting a range of new options in architecting backup and other forms of replication that may reduce data storage costs. Fedora 4 also offers a fixity service. Fixity information typically entails a cryptographic hash (a "checksum") that describes the state of a file at the time of hashing that is widely used as a digital preservation technique to detect if a file has been altered or corrupted.

**Systems Evaluated**

After extensive discussion and analysis, the CMEWG narrowed the focus of further evaluation to three systems: CONTENTdm, Fedora 4/Hydra repository and the LYRASIS Fedora 3/Islandora repository. These three configurations offered the broadest functionality while providing some level of integration with our current architecture.

CONTENTdm is a single software product offered by OCLC that handles the storage, management, and online delivery of digital collections, offering the option of hosting your content locally or on an OCLC-hosted server. Widely used by over 2,000 libraries, CONTENTdm has a more than ten year track record and integrates tightly with other OCLC products, such as Worldshare Management Systems (OCLC's integrated library management system). Currently, the AUC Woodruff Library manages two instances of CONTENTdm: one on behalf of the HBCU Library Alliance, and another which hosts collections that are only available to on-site users due to intellectual property rights. A possible option for the library with regard to continued use of CONTENTdm is the licensing a third instance for online display of AUC Woodruff Library digital collections that are not part of the HBCU Alliance.

Advantages

- Option for cloud based hosting
- Integration with WMS and other OCLC products to enhance discoverability
- In-house familiarity with the product
- Strong tool for access and discovery
- Inline viewer access to audio and video files (via streaming server)

Challenges

- Limits to customization—especially as related to segregated searching, and in branding individual collections
- Ability to manage master files and preservation functionality

Costs

- Licensing
- Software hosting, and-or
- Software support

**Fedora 4 – Hydra**

Fedora 4 with Hydra provides a search-based front end to Fedora 4 that is locally administered. Like Islandora, the Hydra community bridged from Fedora version 3 to version 4. Even in some Fedora 3 implementations of Hydra there are nice features,

such as geo-mapping capacities, which are being extended more formally with Geoblacklight. Fedora 4 implementations now add storage flexibility, potentially offering cost savings in terms of storage space needed. The developer community is typically to be comprised of large research libraries. However, recent developments indicate that open source repository platforms will emerge that can be adopted by a wide range of institutions; in particular, the IMLS funded Hydra-in-a-Box project where DPLA, Stanford, and DuraSpace partner to produce a turnkey, Hydra-based solution intended to appeal to many types and sizes of institutions. (see https://wiki.duraspace.org/display/hydra/Hydra+in+a+Box).

Advantages

- Storage flexibility may enable economical mix of storage
- Transition from versions 3 to 4 appears to be settling now (as of early 2015)
- Responsive web presence for mobile devices
- Potential to add Avalon Media Services Hydra head for searching and online video presentation
- Future out-of-the-box version (IMLS grant)

Challenges

- No direct options for paid hosting; consultants and outside experts are available for development assistance, but there are no direct pay-and-play options for Hydra
- Requires in-house development

Costs

- Developer time
- Possible outside developer assistance
- Ongoing upgrades, maintenance, troubleshooting

**Fedora 3 – Islandora**

Fedora 3 with Islandora provides a front end to Fedora 3 using the open source tool Islandora. Through LYRASIS, institutions can opt for system hosting and support. Software may be locally hosted in conjunction with storage provided by LYRASIS. Islandora works with the Drupal content management system. Originally written in 1999, Drupal gained additional momentum as a

community project by 2005. It is now in wide use as a content management system for building websites. Islandora emerged as an international project via work by University of Prince Edward Island (Australia) to connect Drupal to Fedora.

Advantages

- Can be hosted by LYRASIS or hosted by home institution
- Available as of June 2014
- Involves minimal local development if hosted by LYRASIS

Challenges

- Lacks a 'responsive' web template that adapts to various screen sizes; unclear what mobile presentation would be available
- Likely to remain on Fedora 3 until both Drupal 8 and Fedora 4.2 are available
- Potential remote storage costs for large data footprints (e.g. AV files)
- At review, no streaming AV option without additional development

Costs

- Set-up and migration
- Software hosting
- Software support

**Conclusion**

As a result of efforts by the AUC Woodruff Library's Content Management Evaluation Working Group, we significantly broadened our expertise and knowledge of the digital content management and preservation landscapes. This initiative provided the opportunity to focus on digital content management more comprehensively rather than completing one digital conversion project after the next. The working group considered both short- and long-term goals for the institution as digital content continues to grow and expand. We learned first-hand that this rapidly expanding area of the field requires ongoing monitoring of developments and expansion of skills. After considering a wide range of options, the CMEWG offered a series of recommendations for moving the library

forward in digital content management that reflects the original goals of platform consolidation, interoperability with existing systems, digital asset management, and preservation. The recommendations are in two categories: short-term goals implemented relatively easily with minimal costs, and long-term goals targeted for future consideration and incorporation into strategic and budgetary planning.

Short-Term Recommendations

- Within ArchivesSpace, incorporate the digital object management functionality where possible.
- Work with Archives Research Center staff to evaluate and (ideally) adopt the public access functionality of ArchivesSpace, potentially eliminating a redundancy in a need for XTF.
- Build a Fedora 4-Hydra demonstration repository in-house for testing purposes.
- Develop policies for digital preservation and born-digital collections.
- Continue to expand staff skills and expertise on digital preservation.

Long-Term Recommendations

- Further evaluate Archivematica for preservation functionality.
- Further evaluate MetaArchive for preservation of local content.
- Develop public interface for Fedora-Hydra repository for online display of audio, video, and other digital collections.
- Consider further consolidation and migration of digital collections to Fedora-Hydra repository.

Although there is no final decision about adopting a repository system at the time this was written, staff began implementing the short-term goals outlined by the working group. Digital Services and Archives staff successfully migrated archival management data from Archivists' Toolkit to ArchivesSpace and continues to evaluate the possibility of adopting the public interface. The recent addition of a new Metadata and Digital Resources Librarian position will help

standardize practices and workflows for data creation and management. In addition, discussions commenced about born-digital accessioning and workflows centered on following Society of American Archivists' guidelines to address born-digital materials as outlined in the "Jump-In Initiative," beginning with conducting an inventory of holdings. In 2013, the library became a member of the Digital Library Federation, which has contributed to expanding local knowledge about best practices and emerging trends in digital curation. Through continued internal and external collaborations, monitoring emerging trends, and developing best practices, the AUC Woodruff Library will be well positioned to support the challenges of the long-term stewardship of digital collections.

**Christine Wiseman** has been Head of the Digital Services Unit at the at Atlanta University Center's Robert W. Woodruff Library since 2012. In this capacity she works collaboratively with library staff, the Archives Research Center, and faculties to deliver digital services in support of teaching, learning, and research. She is responsible for implementing best practices and emerging technologies that support the creation, management, discovery, access and preservation of digital collections. Christine serves on the board of the Georgia Archives Institute, is Chair of the Heritage Emergency Response Alliance (HERA), and was the 2011 president of the Society of Georgia Archivists. She lectures on a wide variety of archives and preservation topics, and is a part time instructor for Clayton State University's Masters in Archival Studies program. Christine holds a MLIS and a Certificate of Advanced Study in Preservation Administration from The University of Texas at Austin.

**Al Matthews** is a software engineer. He holds an M.S. from Georgia Institute of Technology and second-authored this publication while working at AUC Robert W. Woodruff Library. For some time he likewise co-chaired Preservation Committee for MetaArchive, a PLN or Private LOCKSS Network. He maintains an independent engineering consultancy as well as a digital arts practice and lives online at http://www.appressorium.com. He lectures in music at

Georgia Tech.

**APPENDIX A: Features and Requirements Checklist for Digital Repository Software**

| **Feature or Requirement** |
|---|
| |
| **General** |
| Replaces current Library CMS (Digital Commons) |
| Replaces current Library CMS (CONTENTdm) |
| Offers integrated management of all digital, electronic collections. |
| Hosted cloud environment based system - an option. |
| Vendor hosted with migrations and data updates to be carried out by the vendor. |
| Support for APIs and/or other interfaces that will allow the library to develop extensions to the core software. |
| Interoperability with a variety of OCLC library resource discovery platforms. |
| Offers multiple options for deposit of digital materials: end user, bulk load, etc. |
| Supports pre-defined workflows for upload of digitized material and their metadata. |
| Supports a variety of metadata standards including but not limited to Dublin Core, etc. |
| |
| **CONTENT MANAGEMENT** |
| Allows for multiple collections w separate branding within same installation of system |
| Allows repository administrator to set submission parameters |
| Home page for each collection |
| Submission Roles |
| Configurable submission roles within collections |
| Email notification for users |
| Email notification for administrators |
| Allows staff to review completed content b/f publication |
| Allows staff to review uncompleted content |
| Allows content administrators to review |

| **Content Acquisition** |
| --- |
| Upload master and access files |
| Batch import of objects/files |
| Batch import of metadata |
| Batch export/content portability (to other systems) |
| |
| **Document/Object Formats** |
| Administrator ability to limit approved file formats |
| Submitted items can comprise multiple files or file types |
| |
| **Text Formats** |
| ASCII |
| UNICODE |
| **Image Formats** |
| TIFF |
| JPEG |
| JPEG-2000 |
| **Presentation Formats** |
| PDF |
| **Audio and Video** |
| WAV |
| Real |
| MP3 |
| AVI |
| MPEG |
| Uncompressed Audio and Video |
| Supports streaming |
| |
| **Version Control** |
| Allow past versions of files to be retrieved |
| Changes identified |
| Changes compared |
| |

| **METADATA** |
| --- |
| Dublin Core |
| EAD |
| METS |
| MODS |
| VRA Core |
| Ability to add/delete customize metadata fields |
| Set default values for metadata |
| Supports import and export (with no loss of data) in all supported formats. |
| Supports PREMIS data model and data dictionary. |
| |
| **SEARCH/USER INTERFACE** |
| Full text |
| Search all descriptive metadata |
| Search selected metadata fields |
| Browse |
| Ability to sort search results |
| Supports integration with library search and discovery tools |
| Viewer for zooming, panning |
| Social media features for commenting, tagging, rating items |
| Support for mobile or responsive themes |
| |
| **ACCESS CONTROL AND PRIVACY** |
| Supports a robust and flexible yet straight-forward system for assigning roles and permissions to staff functions. |
| Supports authorization/authentication which is role/attribute based. |
| Ability to limit access at the collection level |
| Ability to limit access at the file level |
| Ability to define user roles/permissions |
| Ability to integrate with existing security measures |
| |
| **PRESERVATION** |
| Offers persistent document identification |
| Supports PREMIS data model and dictionary |

| |
|---|
| Ability to ingest disk images |
| |
| **REPORTING AND ANALYTICS** |
| Reporting system supports the customization of reports by librarians; this includes but not limited to: changing of reports parameters, views, time range etc. |
| Includes a dashboard in which it is possible to monitor ingest/uploading. |
| Ability to analyze historical data and provide trends analysis. |
| Includes a dashboard in which it is possible to monitor collection usage and downloads |
| |
| **SYSTEM ADMINISTRATION AND MANAGEMENT** |
| Supports basic fulfillment capabilities during local institution network outage. |
| Supports linking of digital resources to the relevant physical/electronic resources in library catalog |
| Comes with "Out of the Box" definitions and configurations so that the library need only make minimal changes to the standard settings. |
| Access to documentation and manuals |
| Customizable to the extent that it can be branded with the library identity. This includes control of style, images and graphical elements, and permits offline stylesheet testing via mockups, development instances, or similar means. |
| Access to mailing list/discussion forum |
| Offers help desk support |
| Offers bug track/feature request system |

**Further Reading**

ACRL Planning and Review Committee. "2012 Top Ten Trends in Academic Libraries," *College and Research Libraries News*, 73 (June 2012), http://crln.acrl.org/content/73/6/311.full.

Bogus, Ian, George Blood, and Robin L. Dale, et. al., *Minimum Digitization Capture Recommendations*. Chicago: Association for Library Collections and Technical Services, Preservation and Reformatting Section, June 2013, http://www.ala.org/alcts/resources/preserv/minimum-digitization-capture-recommendations.

Boyd, Kate Foster and Alma Creighton. "Building a Digital Library on a Shoestring." *Computers in Libraries* 26 no. 6 (June 2006): 14-20.

Chad, Ken. "Unified Library Resource Management Specification." *Open Specifications for Library Systems (LibTechRFP).* Ken Chad Consulting Ltd., 24 July 2012, https://libtechrfp.wikispaces.com/Unified+library+resource+management+specification.

DeRidder, Jody. "*Choosing Software for an Institutional Repository*," April 1, 2004, http://jodyderidder.com/writings/other/ir_software.pdf.

Dion, Hoe-Lian Goh, Alton Chua, Davinia Anqui Khoo, Emily Boon-Hui Koo, et al. "A Checklist for Evaluating Digital Library Software." *Online Information Review* 30 no. 4 (2006): 360-379, doi: 10.1108/14684520610686283.

Dooley, Jackie and Katherine Luce. *Taking Our Pulse: The OCLC Research Survey of Special Collections and Archives.* Dublin, OH: OCLC Research, 2010, http://www.oclc.org/research/publications/library/2010/2010-11.pdf.

Gilbert, Heather and Tyler Mobley, "Breaking Up with CONTENTdm: Why and How One Institution Took the Leap to Open Source."*Code4Lib Journal*, 20 (April 7, 2013), http://journal.code4lib.org/articles/8327.

Kucsma, Jason, Reiss, Kevin, and Angela Sidman "Using Omeka to Build Digital Collections: The METRO Case Study." *D-Lib Magazine*, 16, no. 3-4 (March/April 2010), http://www.dlib.org/dlib/march10/kucsma/03kucsma.html.

Jiang, Lingling, Pham, Kim, Allain, Sara, et al , "Bye Bye, ContentDM! A Migration to Islandora," University of Toronto Scarborough Library, Digital Scholarship Unit, June 2014, http://www.slideshare.net/digitalscholarship/bye-contentdm.

Marill, Jennifer L. and Edward C. Luczak. "Evaluation of Digital Repository Software at the National Library of Medicine." D-Lib Magazine. 15:5/6 (May/June 2009), doi: 10.1045/may2009-marill.

POWRR. "Preserving Digital Objects with Restricted Resources," http://digitalpowrr.niu.edu/tool-grid/.

Van Garderen, Peter and Courtney C. Mumma, "Realizing the Archivematica Vision: Delivering a Comprehensive and Free OAIS Implementation," http://purl.pt/24107/1/iPres2013_PDF/Realizing%20the%20Archivematica%20vision%20delivering%20a%20comprehensive%20and%20free%20OAIS%20implementation.pdf.

Weise, John. "Brief Survey of Digital Library Software Systems," *Library Tech Talk Blog*. University of Michigan, July 8. 2010. http://www.lib.umich.edu/blogs/library-tech-talk/brief-survey-digital-library-software-systems.

Yale University. "Hydra/Fedora Presentation," August 2014, http://web.library.yale.edu/sites/default/files/files/HydraFedoraMSSA_presentation-v2.pdf.